Journal of Financial Economics 113 (2014) 131–155

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Performance evaluation with high moments and disaster risk ☆

Ohad Kadan *, Fang Liu

Olin Business School, Washington University in St. Louis, United States

A R T I C L E I N F O

Article history:
Received 2 April 2012
Received in revised form
4 June 2013
Accepted 20 October 2013
Available online 19 March 2014

JEL classification:
D81
G11

Keywords:
Performance evaluation
Rare disasters
High distribution moments

A B S T R A C T

Traditional performance evaluation measures do not account for tail events and rare disasters. To address this issue, we reinterpret the riskiness measures of Aumann and Serrano (2008) and Foster and Hart (2009) as performance indices. We derive the moment properties of these indices and their sensitivity to rare disasters and show that they are consistent with the asset pricing literature. As applications, we show that "anomalous" investment strategies such as "momentum" or investment in private equity lose much of their glamour when accounting for high moments and rare events. Furthermore, using the indices to select mutual funds results in desirable high-moment properties out of sample.

1. Introduction

Tail risk and rare disasters have been central to the recent meltdown in financial markets. Indeed, markets were hit by catastrophic events whose ex ante probabilities were considered negligible. Traditional performance evaluation measures (such as the Sharpe ratio) typically rely on the first two distribution moments, thereby underestimating the effects of rare disasters. Indeed, low distribution moments hardly account for rare and catastrophic events, since their large negative effect is multiplied by a very small probability. By contrast, when one considers high distribution moments, an extremely negative but rare outcome is raised to a high power, making its effect on the moment substantial regardless of the small probability associated with it.

High distribution moments have received notable attention in the asset pricing literature. In particular, a large body of work in asset pricing suggests that investors favor right skewness (e.g., Rubinstein, 1973; Kraus and Litzenberger, 1976; Jean, 1971; Kane, 1982; Harvey and Siddique, 2000), but are averse to tail-risk and rare disasters (e.g., Barro, 2006, 2009; Gabaix, 2008, 2012; Gourio, 2012; Chen, Joslin, and Tran, 2012; Wachter, 2013). It is thus desirable that normative performance evaluation measures reflect these preferences.

In this paper we study two such performance indices relying on a simple reinterpretation of the novel riskiness measures proposed by Aumann and Serrano (2008) and Foster and Hart (2009) (hereafter AS and FH, respectively).[1] We investigate the moment properties of these indices and establish that they reflect *all* distribution moments in a

☆ We thank an anonymous referee, Phil Dybvig, Sergiu Hart, Isaac Kleshchelski, Asaf Manela, Mark Rubinstein, and Shlomo Yitzhaki as well as seminar participants at the University of Arizona, Interdisciplinary Center, Hertzlia, the Institute of Financial Studies at Southwestern University of Finance and Economics, China, and Washington University in St. Louis for helpful comments and suggestions.

* Corresponding author.
E-mail addresses: kadan@wustl.edu (O. Kadan), fliu23@wustl.edu (F. Liu).

[1] Aumann and Serrano (2008) offer a set of axioms characterizing the AS riskiness measure. An axiomatization of the FH measure is offered separately in Foster and Hart (2013).

http://dx.doi.org/10.1016/j.jfineco.2014.03.006

manner consistent with economic intuition and with the asset pricing literature. We also discuss the way these two indices reflect disaster risk. We then apply these indices to popular investment strategies and to well-known anomalies, show their practical usefulness in selecting mutual funds, and demonstrate the pitfalls associated with ignoring high moments and rare disasters in performance evaluation.

Our starting point is that investors are risk-averse and choose their investments by maximizing expected utility. The best possible way to rank investments in this setup is known to be Second Order Stochastic Dominance (SOSD) (see Hadar and Russell, 1969; Hanoch and Levy, 1969; Rothschild and Stiglitz, 1970), according to which one investment dominates another if all risk-averse investors prefer the former to the latter. The problem with SOSD is that it only imposes a partial order on investments. Namely, some pairs of investments cannot be ranked using SOSD.

Based on our discussion thus far, a desirable performance evaluation index should satisfy the following four requirements: (i) impose a complete order on investments, namely, any two investments can be compared; (ii) depend on the distribution of outcomes only. That is, the form of the utility function is not needed to calculate the performance index; (iii) coincide with SOSD, whenever SOSD can be applied. Namely, if all risk-averse investors prefer one investment to the other, then the performance index ranks the investments accordingly; and (iv) account for high distribution moments in a manner consistent with the asset pricing literature. That is, the index is increasing in mean and skewness and decreasing in variance and tail-risk of the investment.

The Sharpe ratio, which is probably the most popular performance evaluation measure, satisfies (i) and (ii), but clearly fails (iv). Interestingly, it also fails (iii). Indeed, it is fairly easy to find examples in which all risk-averse investors prefer one investment to the other and yet the Sharpe ratio ranks the investments in the wrong order (see Section 2 for examples). In Appendix B we review several other popular performance evaluation measures and discuss the extent to which they satisfy these four requirements (Table B1).

To understand the fundamental insights in AS and FH it is useful to follow the approach presented in Hart (2011), who offers a unified framework for the two. The key for the new indices is to use the investor's initial wealth as a benchmark for her investment decisions. That is, instead of comparing the expected utility of two investments, we compare the expected utility of each investment separately to the status quo, and ask which one of the two investments is uniformly rejected more often. If each time that investment g is uniformly rejected we have that investment g' is also uniformly rejected, then g is deemed more attractive than g' (i.e., g has better performance than g').[2] That is, g is more attractive than g' if g is rejected "less often" than g' in some uniform manner when compared to the status quo.

The term "uniform rejection" can take two different meanings. First is "wealth-uniform rejection" in which for a given utility function, an investor rejects the investment relative to the status quo for all wealth levels. Second is "utility-uniform rejection" in which for a given wealth level, all utility functions reject the investment relative to the status quo. The former approach to uniform rejection leads to the AS performance index, while the latter leads to the FH performance index.

As shown in AS, FH, and Hart (2011), the two approaches yield two rankings of investments, each of which can be represented by a positive performance index that possesses an intuitive economic interpretation. Both indices satisfy requirements (i)–(iii) above. Moreover, they can be easily calculated from the distribution of the investment by solving an intuitive implicit equation. The only difference between our interpretation and the interpretations in AS and FH is that they choose to consider the riskiness of the investment, deeming one investment "*more risky*" than another if it is uniformly rejected *more* often relative to the status quo. We choose to focus on the flip side of the argument, viewing one investment as "*more attractive*" than another if it is uniformly rejected *less* often relative to the status quo. Roughly speaking, we view an investment as "attractive" if risk-averse investors show little aversion to this investment when compared to the status quo, in a uniform manner.

The first thing we do in this paper is to extend the AS and FH indices to a multi-period setting. We show that the AS and FH results can readily be considered in such a setting, and that if gambles are identically distributed in each period, then the multi-period performance indices coincide with the single-period indices.

We then turn to studying how the AS and FH performance indices are affected by the moments of the investments being evaluated. We establish that both the AS and FH indices reflect *all* the distribution moments (raw and central). Moreover, these performance indices are increasing in all odd moments and decreasing in all even moments. Consequently, the two indices satisfy requirement (iv) above.

Next, we ask whether the sensitivity of the performance indices to the moments is monotonically decreasing in the order of the moment. Namely, do high distribution moments necessarily have a smaller effect on performance than low distribution moments? We establish that there is no such monotone relation. In particular, the performance indices can be either more or less sensitive to higher moments. Thus, high moments can have a material effect on performance, and should not be neglected.

We then turn to exploring how the performance indices are affected by rare disasters, modeled as extremely negative outcomes associated with vanishing probabilities. First, note that such outcomes tend to make the distribution left skewed (more negative third moment) and fat-tailed (higher fourth moment). Thus, given requirement (iv), both performance indices are adversely affected by rare disasters. However, we show that the FH

[2] The term "investment" here simply refers to a random variable which can be described by the probability distribution over outcomes. We often use the term "gamble," which is the one used in AS and FH, instead. We use the letter g as a generic notation for such investments (or gambles).